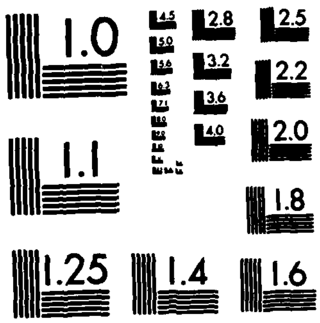
AD-A144 067 VALIDATION OF RELATIVE-TIME-SPENT RATING SCALES(U) AIR FORCE HUMAN RESOURCES LAB BROOKS AFB TX S K GARCIA JUL 84 AFHRL-TP-84-11 1/1

UNCLASSIFIED F/G 5/9 NL

END



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFHRL-TP-84-11

4

AIR FORCE

HUMAN RESOURCES

LABORATORY

VALIDATION OF RELATIVE-TIME-SPENT RATING SCALES

By

Sharon K. Garcia

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235

July 1984

Interim Paper for Period July 1981 – December 1981

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Public Affairs Office has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

NANCY GUINN, Technical Director
Manpower and Personnel Division

J.P. AMOR, Colonel, USAF
Chief, Manpower and Personnel Division

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

AD-A144 067

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFHRL-TP-84-11	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Manpower and Personnel Division Air Force Human Resources Laboratory 6b. OFFICE SYMBOL (If applicable) AFHRL/MOD	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code) Brooks Air Force Base, Texas 78235	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Human Resources Laboratory 8b. OFFICE SYMBOL (If applicable) HQ AFHRL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code) Brooks Air Force Base, Texas 78235	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 62703F PROJECT NO.: 7719 TASK NO.: 19 WORK UNIT NO.: 11

11. TITLE (Include Security Classification)
Validation of Relative-Time-Spent Rating Scales

12. PERSONAL AUTHOR(S)
Garcia, Sharon K.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Interim	FROM Jul 81 TO Dec 81	July 1984	42

16. SUPPLEMENTARY NOTATION
Work performed under TS Study Number 7849

17. COSATI CODES

FIELD	GROUP	SUB. GR.
05	09	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
absolute time
CODAP
job inventory
occupational analysis
rating scales
scale validation
time estimation

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Relative-time-spent rating scales are used as the primary measuring device in task-oriented job inventories. These scales permit incumbents to report the amount of work time they spend on each task performed relative to time spent on other tasks. Measures of relative time spent are currently being collected by the Air Force and other governmental agencies; however, no consensus has been reached regarding the optimal scale format to use in obtaining time-spent-performing data. The general lack of consensus regarding the optimal scale has stemmed primarily from the differences among scientists in their opinions about scaling procedures, scale construction, application of scales and validity of scales. This paper summarizes the results of a feasibility study conducted to validate various relative-time-spent scale formats. The criterion for validation was collected via direct field observations. The primary objective of this investigation was to determine the relative validity of binary (perform/not perform), 9- and 25-point scales using actual time spent and frequency of observed task performance criteria. Results of this investigation indicated that the 9-point relative-time-spent scale provided the optimal format for use in the Air Force occupational analysis program.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Including Area Code)	22c. OFFICE SYMBOL
Nancy A. Perrigo Chief, STINFO Office	(512) 536-3877	AFHRL/TSR

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

AFHRL Technical Paper 84-11 July 1984

VALIDATION OF RELATIVE-TIME-SPENT RATING SCALES

By

Sharon K. Garcia

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235

Reviewed by

Hendrick W. Ruck
Chief, Job Requirements and Training Function

Submitted for publication by

R. Bruce Gould
Chief, Force Utilization Branch

This publication is primarily a working paper.
It is published solely to document work performed.

PREFACE

This effort was completed under Work Unit Number 77191911, "Measurement and Analysis of Job and Mission Requirements."

The author wishes to express appreciation to TSgt William Anderson, the Technical Services Division project officer for this study, to Sgt Louis M. Datko, SrA Grace R. Hrymoc, A1C Jack Nelson, and A1C Ann Sernak who served as the project observers, and to Mrs Connie Villareal for her programming support. Finally, special thanks is expressed to the airmen and their supervisors who participated in support of this effort.

Special appreciation is extended to Dr. Hendrick W. Ruck and Dr. William E. Alley for their technical guidance.

TABLE OF CONTENTS

List of Figures

List of Tables

VALIDATION OF RELATIVE-TIME-SPENT RATING SCALES

I. INTRODUCTION

Occupational analysis in the United States Air Force is primarily concerned with obtaining viable and quantifiable information about jobs performed by Air Force personnel. This information is obtained through analyses of work performed, of characteristics of job incumbents, and of the occupational structure. Information derived through occupational analysis is fundamental to personnel management and is essential for personnel selection, classification, job proficiency measurement, and training curricula development.

The Air Force occupational analysis program uses a job inventory method to obtain information about work performed. Briefly, this method employs a survey booklet (job inventory) which lists tasks that are performed by persons within an Air Force specialty (AFS) or occupation. In completing a job inventory, individuals check each task they perform as a part of their regular job. They then rate the tasks which they perform using a 9-point relative-time-spent rating scale. A rating of "1" indicates that the individual spends very little time on a task in relation to other tasks he or she performs. A rating of "9" indicates that a very large amount of time is spent in the performance of that task compared to other tasks.

The relative-time-spent rating scale is the primary measuring device used with job inventories to gather estimates of the amount of time each worker spends on each task. Relative ratings derived from worker responses to job inventories are converted to estimated percent time spent values. These values indicate a percentage estimate of an individual's work time spent on each task performed. To perform the conversion to relative time, it is assumed that the total of an individual's raw relative-time-spent ratings represents 100% of his or her work time. This assumption is defensible since job inventories are exhaustive lists of tasks that may be performed. Therefore, each raw rating is expressed as a percentage of the total raw ratings across all tasks for that individual. The average time-spent value for any group of workers indicates the percentage of total group time spent on each task. The sum of these group values also equals 100% (Archer, 1966).

The conversion of raw ratings to percentages and the reporting of resulting job descriptions are performed using the Comprehensive Occupational Data Analysis Programs (CODAP), a series of computer programs developed specifically to analyze job inventory data and to describe the work performed by individuals and groups of individuals (Christal, 1974). Individual and group job descriptions can be computed at the task level to show how work time is distributed over the tasks performed. Relative-time-spent values provide a convenient method for computing the overlap (or similarity) (a) of two individual jobs with each other, (b) of an individual job with a group job description, or (c) of one group job description with another group job description.

Historically, time-spent rating scales have varied greatly. Characteristics that have varied include scale length (number of points), anchoring of scale points, and scale interval used. This investigation will be limited to issues associated with scale length.

II. BACKGROUND

A major issue in this effort is the criterion; hence, this section will include a short review of specific criterion measures.

Within the Air Force Occupational Analysis Program (Christal, 1974), a 9-point relative-time-spent scale is used for collecting time-spent performing tasks data. Prior to the use of the 9-point scale, a 7-point relative-time-spent scale was used operationally by the Air Force. Christal and Madden (1961) recommended the use of a 9-point scale, arguing that a larger number of scale points may be needed to ensure that "extreme" jobs or jobs which are "off scale" could be appropriately rated. The issue raised by Christal and Madden paved the way for a series of investigative studies designed to address the optimal number of scale points to use in job measurement.

Scale Length Comparisons

Comparisons of 5-, 7-, 9-, 25-, and 100-Point Scales

Carpenter, Giorgia, and McFarland (1975) conducted two studies to address the issue of optimal number of points by investigating the effect of potentially useable scale formats on the accuracy of the derived job description. Accuracy was operationally defined as the difference between job descriptions based on the sum of the absolute values of time spent on tasks (Archer, 1966).

In the first study, the experimental stimuli were five hypothetical job descriptions. These job descriptions specified the tasks performed and the average time in minutes spent on each task per week. Subjects were 265 airmen randomly assigned to one of five groups (53 per group). Each airman was given a job description which he or she was to assume to be an accurate representation of a job he or she was performing. Each subject then completed job inventories using 5-,7-,9-, and 25-point scales and a scale in which they directly estimated the percent of time spent.

Since the actual time spent per week on each task was available, accurate percentage values were computed to serve as the criterion against which the derived job descriptions could be evaluated. The percent time-spent values derived from the relative scales, and the direct estimate of percent time spent, were then compared to the criterion values and the absolute percentage differences across all tasks were summed to derive a criterion comparison (CRICOM) value. The CRICOM was used as the dependent variable in a two-way analysis of variance with repeated measures on the scale effect variable. Results of the analysis indicated that the 5-point rating scale was significantly inferior in accuracy to all other scales

employed and that the 7- and 25-point scales were inferior to the direct estimate of percent time spent. Of all the scales used in this study, the 9-point scale was found to be the statistically more accurate.

In the second study (Carpenter et al., 1975), a job inventory was constructed which included all tasks performed by Air Force basic military trainees. Eighteen flights of basic airmen (N = 834) were identified as subjects. Basic training instructors were provided with copies of the task inventories and recorded in minutes the actual time spent on each task performed by members of their flight. The basic trainees were randomly divided into eight groups. Each group used one of eight scales for the derivation of individual job descriptions. Five of these scales measured relative-time-spent and differed in terms of the number of points. The scales were 5-, 7-, 9- (all scale points anchored), 9- (extreme and midpoint anchored) and 25-point scales. The remaining scales measured relative frequency of task performance, estimates of actual time on tasks, and actual percent time spent on tasks.

Individual percent time spent descriptions were derived ar[illegible]RICOM values as previously defined, were computed. These values [illegible]e then analyzed, using a one-way analysis of variance, to determine th[illegible]elative validities for subjective time-rating scales. The 9-point scale, [illegible]ardless of the number of anchors provided, was again found to be statist[illegible] more accurate than the other scales used. This conclusion was based c[illegible] lower obtained CRICOM value for the two 9-point scales as compared to that for the other scales used. From these results, the 9-point relative time spent scale was adopted as the primary scale for use in collecting time spent performing data in the Air Force.

Comparison of Dichotomous Scale to 5-Point Scale

A study which reached somewhat different conclusions was conducted at the Navy Personnel Research and Development Center by Pass and Robertson (1978). Of primary interest in that effort were the stability and interrelationship of two types of job task scales, namely the relative-time-spent scale (5-point) and the dichotomous task-performed scale. Scale stability was evaluated by comparing the job description profiles of average scale scores between randomly split samples for each pay grade within four occupations. Scale interrelationship was evaluated by comparing the job description profiles across scales. Results of this investigation indicated that the stability of both the continuous and dichotomous scales was quite high. When the average relative-time-spent per task was calculated for only those incumbents actually performing a task, the stability was very low. This is understandable because a large amount of variance is expected across job incumbents in time spent on tasks. In addition, the average value for each task on the relative-time-spent scale was generally very small; often less than 1% of the total time spent, suggesting that members in a pay grade spend, on the average, less than 1% of their time performing any single task. This is not surprising given that the average worker performs more than 100 tasks.

Based on these results, Pass and Robertson (1978) suggested that the relative-time-spent scale be deleted from task inventories. In its place, they recommended that a dichotomous scale be substituted. It must be noted that their study did not include a validation of either scale on any external criteria.

In response to this recommendation, Carpenter and Giorgia (1979) reanalyzed data obtained in their previous study (Carpenter et al., 1975) converting the original scale length formats into perform/not perform (1/0) data. Individual job descriptions were then derived based on these new data. The magnitude of the CRICOM variance was found to be significantly greater than that reported for any of the relative scale formats.

The results of the Carpenter and Giorgia (1979) study indicated that, for the group of subjects using the 5-point rating scale, the CRICOM values were smaller than the CRICOM values computed from their dichotomized responses. Hence, the relative-time-spent scale possessed significantly greater inherent accuracy than the dichotomous scale in terms of the subject's ability to provide an accurate portrayal of his or her job. Similar analyses were computed using those subjects who had originally employed the 9- and 25-point scales. Results again favored the use of the relative-time-spent scale.

The results clearly indicated that the accuracy of the derived job description was functionally related to the number of scale points available to the respondent. Further, the consistency of the 1979 results with the 1975 research findings suggested that the use of a perform/not perform response on the part of the incumbents, although highly stable, may result in individual job descriptions that are significantly inferior to those based on scales providing a wider range of response options. It was concluded that if the validity of an individual job description is considered important for personnel actions or decisions, then the scale format providing the optimal accuracy is to be desired (Carpenter et al., 1975). They found the 9-point scale format to be the most accurate format.

Scale Point Evaluations

Number of Scale Points Used

Following the lead of Carpenter et al. (1975), Wissman (1980) examined the relationship between scale length and the number of points used. In his study, Wissman evaluated six scale lengths. In particular, he examined 5-, 7-, 9-("all" points anchored and "end and middle" points anchored), 25-, and 99-point scales. A sample of 1,800 first-term airmen within the Fuels career ladder was used. Each subject was instructed to complete a job inventory using one of the scales provided. Wissman found that for the 5- and 7-point scales, the average number of points used was four and five, respectively. The number of points used increased to six for the two variations of the 9-point scale as well as for the 99-point scale and seven for the 25-point scale. Thus, based on Wissman's analyses, the average number of points used by individuals completing job inventories never

exceeded seven points, regardless of scale length. While these results are suggestive, they do not provide strong evidence in favor of any one scale.

Direct Observation as a Criterion

Perhaps the major difficulty in determining the best scale to use is the determination of and agreement among researchers regarding a valid criterion on which to evaluate a scale's accuracy and efficiency.

In an attempt to address the issue of validity, Hartley et al. (1977), incorporated objective methods in evaluating validity of self reports. Their study used an observation method to provide criterion data for assessing the accuracy of three types of self-report information. The three types were (a) the identification of which task had been performed, (b) the relative amount of time spent per task, and (c) the amount of time spent on any one task. Work behavior was sampled every 30 seconds for one day per subject to provide criterion data. Results confirmed the Air Force findings cited by McCormick (1976) that relative time judgments by workers are better than absolute time judgments. However, Hartley et al. (1977), stated that when it is critical to have precise information about the worker's job and the time he or she spends doing various tasks, an objective observational approach may be more appropriate than worker judgments.

In a similar study, Johnson, Hiller, and Tokunaga (1980) attempted a validation of a job analysis questionnaire using observational data as the criterion. The correlation between the questionnaire data and observational data was significant (r = .86, $p \leq .01$), indicating that an observational approach to validating questionnaire data was feasible. According to Johnson et al. (1980), the data collected by job inventories were found to be reliably related to the observation criterion. Hartley et al. (1977) and Johnson et al. (1980) showed that relative-time-spent scales could be validated using on-site observations as criteria; however, neither study addressed the issue of scale length.

III. OBJECTIVES

The present investigation was conducted to examine the feasibility of using direct field observation data to validate relative-time-spent rating scales of varying lengths. The primary objectives of this effort were to compare binary (perform/not perform), 9- and 25-point time spent scales in terms of (a) their characteristic response patterns, (b) their predictive validity with regard to direct field observation, and (c) their stability with different sample sizes. The criteria used were actual time spent and frequency of observed task performance based on a direct observational approach.

IV. METHOD

Subjects

The sample consisted of 21 Computer Programming Specialists (Air Force Specialty Code 511X1), assigned to the Technical Services Division, of the Air Force Human Resources Laboratory at Brooks AFB. Of the total sample, six were apprentice (3-skill) level and 15 were journeyman (5-skill) level enlisted personnel. The sample size was kept small because the feasibility of collecting direct observational data in an Air Force setting was untested.

Measures of Time Spent

Perform/not perform, 9- and 25-point scales were used to operationally represent estimates of actual time spent and will be referred to as the predictor variables throughout the remainder of this paper. The selection of these particular scales for use in this investigation was based on (a) the use of the 9-point scale in the operational Air Force occupational analysis program, (b) the potential ability of the 25-point scale to capture more definitive information about an occupation through the availability of greater response options and (c) the findings (Pass & Robertson, 1978) that suggested the use of a binary (perform/not perform) scale.

Absolute time estimates and frequency of task occurrence based on work sample observations were used as the criteria for evaluating the relative-time-spent scales in terms of their ability to provide the best estimates of observed time spent on tasks.

Procedure

Phase I

The effort was conducted in two phases. Phase I produced the predictor variable data and Phase II, the criteria data. Phase I consisted of the administration of job inventory booklets consisting of 577 tasks to each of the 21 subjects. Each subject completed two job inventories: one with a 9-point relative-time-spent scale having all scale points verbally anchored, and one with a 25-point relative-time-spent scale having only extreme and midpoints anchored (Table 1). The inventories were administered in a counterbalanced design. Eleven subjects completed the inventory with the 9-point relative time spent scale followed by the 25-point inventory. The other 10 completed the 25-point inventory followed by the 9-point. The administration of inventories occurred over a 2-day period with about 24 hours between administrations so as to minimize time effects.

For purposes of obtaining perform/not perform data, a binary coding of task responses was performed for one of the two inventories completed by each individual. The 9-point responses by 11 subjects and the 25-point responses by the other 10 subjects were randomly selected and used for deriving perform/not perform data. A code of "1" was applied to each task checked by the respondent to indicate he or she performs the task, and a "0"

was assigned to those tasks not rated, indicating that he or she does not perform the task.

Table 1. Relative Time Spent Scales

1-9	1-25
1 = very small amount	1 = almost none
2 = much below average	2 =
3 = below average	3 =
4 = slightly below average	4 =
5 = about average	5 =
6 = slightly above average	6 =
7 = above average	7 =
8 = much above average	8 =
9 = very much amount of time	9 =
	10 =
	11 =
	12 =
	13 = average
	14 =
	15 =
	16 =
	17 =
	18 =
	19 =
	20 =
	21 =
	22 =
	23 =
	24 =
	25 = almost all

Note: The perform/not perform scale was obtained from a binary coding of tasks identified as performed by incumbents based on their responses to the 9- and 25-point scales.

Phase II

Phase II involved the collection of direct time estimates based on the tasks actually performed by personnel previously surveyed. Actual time estimates were obtained by direct observation. The direct observations were conducted over a period of 6 weeks by a behavioral scientist and two scientific assistants. These personnel will be referred to as "observers" for the remainder of this paper.

Assignment of Subjects

Observers monitored the activities of seven subjects each day over a 6-week period. Each week a different group of seven subjects was assigned to each observer so that each observer had the opportunity to observe all 21 subjects a number of times over the period.

Observation Schedules

Observers were provided with weekly log sheets for each subject for each of the 6 weeks. The log sheets listed the subject's identification number (1-21), the date of the observation, and a daily goal number of observations (four or six) for each subject per week. The goal numbers were established to achieve a total of 130 observations per subject over the 6-week observation period.

Observation times were randomly assigned for each subject for each day throughout the period. Times were established to assure that tasks performed would be measured at different times throughout the day during the observation period and to minimize the conflicts in observer time schedules. Observation times were divided into 5-minute increments (i.e., 0805, 0915, etc.). The times ranged from 0715 to 1600 hours, with a break in observations between 1100 and 1230 hours.

Observation Procedures

At the scheduled times, the observer located his or her assigned subject and recorded information relating to the following:

1. The task currently being performed

2. The time at which that task was begun

3. The estimated time of completion for that task

In the event an observer had difficulty determining the task being performed by observation (i.e., a program run in batch versus demand mode), the subject was queried to assure accuracy in recording the task being performed. Information provided by the subject was then recorded by the observer in the appropriate columns of the log sheet. The observer's daily log sheet consisted of six columns (Figure 1). At the top of each sheet, a space was provided to identify the subject being observed, the day and date of the observation, and the observer or monitor. Observers recorded the following specific information:

1. Observation number: the daily goal number of the observations being made for a subject.

2. Observation time: the actual time at which an observation was made.

Week 1

Subject: Monitor:

Date	Observation Number	Observation Time	Task No.	Start Time (S)	End Time (E)	Difference

Figure 1: Sample Daily Log Sheet

3. Task number: job inventory number associated with the task observed at the specified time.

4. Start time (S): subject's estimated time that the observed task was begun.

5. Estimated end time (E): subject's estimated completion time for the task observed.

6. Difference (E-S): total time for observed task as defined by taking the difference between the estimated end time and start time.

In recording the information in the difference column (E-S), the observer subtracted the estimated end time for an observed task from the starting time of that task for each observation to yield a time estimate for each task observed, recorded in minutes (see Table 2).

Table 2. Sample Recording of Data for Varied Task Performance

Observation No.	Observation Time	Task	Start	End	Difference
1	0815	9	0800	0830	30
2	0830	47	0830	0900	30
3	0915	59	0930	0945	15
4	1430	17	1415	1500	45

In the event that performance on a single task was observed more than once in the course of performance such that the start time remained the same but estimated completion times changed over observations, the difference (E-S) was defined as the difference between the final recorded estimate of completion minus the initial start time (see Table 3).

Table 3. Sample Recording of Data for Same Task Performance

Observation No.	Observation Time	Task	Start	End	Difference
1	0745	98	0730	0800	
2	0815	98	0730	0830	
3	0900	98	0730	0930	
4	1000	98	0730	1030	180

In this case, the total time for task 98 was the difference between the start time (0730) and the last recorded entry for the estimated completion time (1030). No entry was made in the difference column for the previous observations of the same task. Nevertheless, each observation of that task was counted as an independent observation for determining the frequency of observed task performance, defined as the total number of times a task was observed as performed. Absolute time on the other hand was defined as the total time taken to complete a task. This rule was applicable only in the event that the same task was observed as being performed with the initial time of start remaining the same over observations for a particular day. In the event that the task was observed as performed consecutively throughout the day with different start and end times for each occurrence, the E-S was defined as outlined in Table 2.

This procedure was followed by each of the observers throughout the 6-week observation period so as to allow for 130 observations for each subject.

Obtained Measures

Information obtained from Phases I and II provided data on the three predictors and two criterion variables for each of the 21 subjects:

1. Relative time spent based on a binary (perform/not perform) scale.
2. Relative time spent based on a 9-point relative time spent scale.
3. Relative time spent based on a 25-point relative time spent scale.
4. Actual time spent converted to relative time (criterion).
5. Frequency of observed task performance converted to relative frequency (criterion).

Relative time spent for each of the five measures was computed by summing all of a subject's Phase I survey responses. Each entry (raw rating, absolute time, and task frequency) for each subject was then divided by the total for that subject, and the quotient was multiplied by 100 to give the percentage of time spent on each task for a given subject. The total for each subject was taken to represent 100% of the subject's work time.

Statistical Analyses

The analyses were designed to compare binary (perform/not perform), 9- and 25-point time-spent scales on their characteristic response patterns, their predictive validity with regard to the criteria, and their stability with different sample sizes.

Descriptive statistics were used to examine the characteristics of the predictor and criterion variables. The total number of points used by each subject and the average number of points used for all subjects across the 577 tasks in the inventory were computed. These values were used to examine the response patterns on the perform/not perform, 9- and 25-point scales.

Two analyses were performed to assess the validity of the various scales: overlap and correlation. Overlap of time spent was calculated between predictor and criterion variables to determine consistency and/or inconsistency between ratings across scales. A second estimate of validity was produced by computing correlations among the three predictors and two criteria.

An issue of particular interest in this investigation was the adequacy of the sample size. To gain insight into this issue, interrater reliability indices were examined as a function of sample size, for each of the five task response scales. Using the CODAP program REXALL (Christal & Weissmuller, 1976), reliability values were first computed for a random sample of only two cases. Additional randomly selected cases were then added, one by one, to the sample, with new reliability values being computed for each scale at each successive stage. Evidence that the reliability values were stabilizing at some point, would suggest the attainment of an adequate sample size.

V. RESULTS

Characteristic Response Patterns

Descriptive Data

Summary statistics for the three predictors and two criteria for the total sample are shown in Table 4. Comparisons among the five measures provide some interesting differences in terms of the number of tasks rated as well as the mean time spent vector. Of the 577 tasks listed in the job inventory, the average number of tasks rated was 60.4 and 61.6 for the 9- and 25-point scales, respectively.

Table 4. Summary Statistics for Predictor/Criterion Variables (N=21)

VARIABLE	AVERAGE NO. OF TASKS RATED	MEAN OF ALL TASK RATINGS	STANDARD DEVIATION
<u>PREDICTOR VARIABLES</u>			
PERFORM/NOT PERFORM SCALE	57.8	.10[a]	.31
9-POINT SCALE	60.4	3.82[a]	2.28
25-POINT SCALE	61.6	9.71[a]	6.52
<u>CRITERION VARIABLES</u>			
TIME	29.3	35.37[b]	25.64
FREQUENCY	29.3	4.06	5.58

[a] Expressed in scale points.
[b] Expressed in minutes.

For the 9-point scale, the mean of all task ratings for the total sample was 3.82. The standard deviation was 2.28. The mean and the standard deviation for the 25-point scale were 9.71 and 6.52, respectively. The average number of tasks performed for the perform/not perform scale was 57.8. The mean of the perform/not perform scale was .10 with a standard deviation of .31.

For the time and frequency criteria, the average number of tasks performed was 29.3 for both variables, indicating a decrease in the average number of tasks rated in relation to the 9- and 25-point scales; however, it should be noted that the the observational approach captured only those tasks observed as performed within a 6-week period, whereas the job inventory approach captures past as well as present tasks performed on the job.

The mean rating for time shows an average time spent of 35.37 minutes per task across the 6-week observation period, with a standard deviation of 25.64. For frequency, the mean reflects the average number of times any one task was observed to be performed across observations. The mean rating for frequency was 4.06 times per task per subject, with a standard deviation of 5.58.

Frequency Data

The distributions of points used for the three predictors are represented in Figures 2 through 4. Figure 2 displays the perform/not perform responses. Tasks performed, as indicated by "1" accounted for 10.7% of the responses, while "0" was used 89.3% of the time. For the 9-point scale (Figure 3), a rating of 2 was used most frequently (1.8%), followed by a rating of 1 (1.7%). A zero point was added to the 9- and 25-point scales to show the percentage of tasks not performed in relation to those that were. For the 25-point scale (Figure 4), scale points 13, 5, 10, and then 20 were used more frequently. This is an interesting pattern, indicating that the most frequently used scale points occurred at approximately 5-point intervals. The fact that 13 was the most frequently used scale point is not suprising since it is one of three points which is anchored and is labeled "average." Had 15 been the midpoint, it is assumed that it would have been the most frequently used point of the scale. For the 9- and 25-point scales, the data show that over all subjects and all tasks, all scale points were utilized.

Number of Scale Points Used

Figure 5 displays the number of scale points used by each subject on the 9- and 25-point scales. For the 9-point scale, the minimum number of points used by any subject was five and the maximum was nine. The average number of points used across all subjects was seven. The minimum number of points used by any subject for the 25-point scale was seven, with a maximum of 23. For the 25-point scale, the average number of points used was 13.

Stability of Scales by Sample Size

To test for the adequacy of the sample size for the three predictors and two criteria, interrater reliability indices (R_{11}) for groups of raters of different sample sizes were derived. This was done by randomly adding a case incrementally to increase the sample from two subjects to the total sample of 21 subjects. The results of this process and the corresponding R_{11} values are plotted on Figure 6. The point at which all five scales had little or no change in reliability was with a sample size of around 15. A general "leveling off" of reliability began to occur at this stage for each of the predictors and criteria. This indicates that for the purposes of this investigation, a sample size of 21 subjects was sufficient to obtain meaningful results.

Validity

To address the issue of validity, two approaches were taken: overlap of time-spent job descriptions of the total sample for each of the five measures was computed, and intercorrelations among predictor and criterion variables were computed. Overlap among the five measures on time spent was computed using the CODAP program OVRLAP (Archer, 1966). The degree to which each scale overlapped with one another in terms of time spent is presented in Table 5. The highest overlap value obtained between scales was found between the 9- and

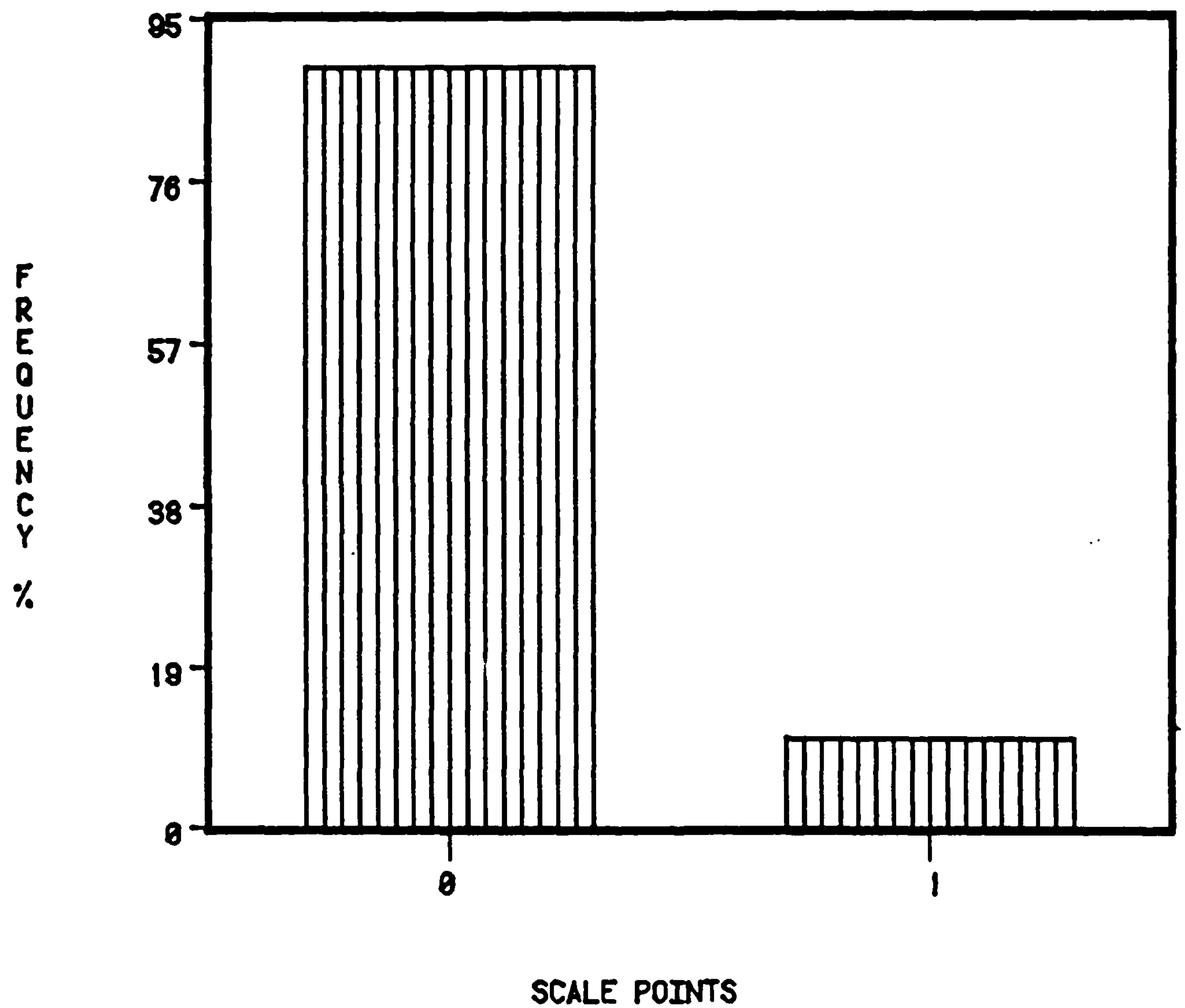


FIGURE 2: FREQUENCY OF POINTS USED FOR THE PERFORM/NOT PERFORM SCALE

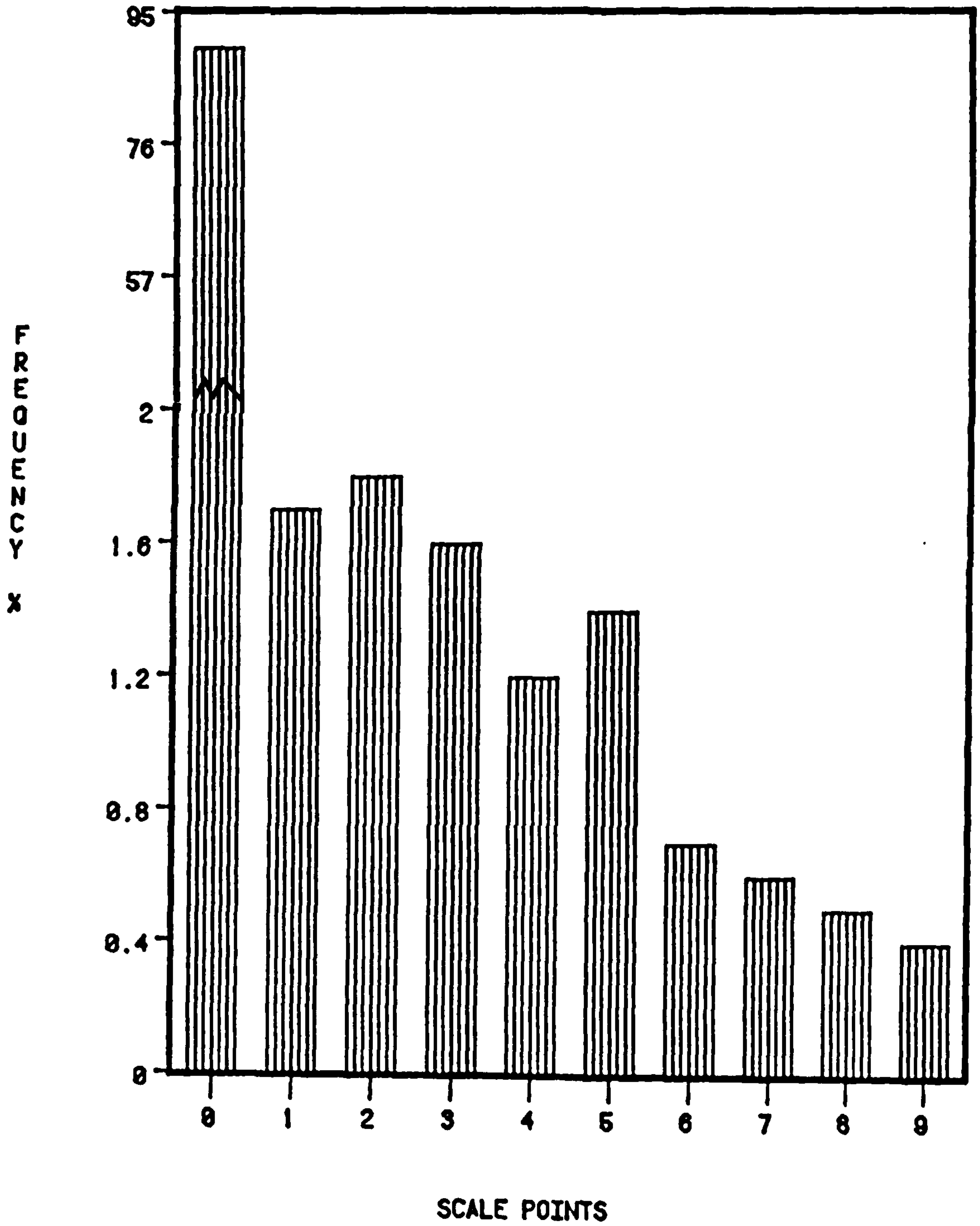


FIGURE 3: FREQUENCY OF POINTS USED FOR THE 9-POINT SCALE

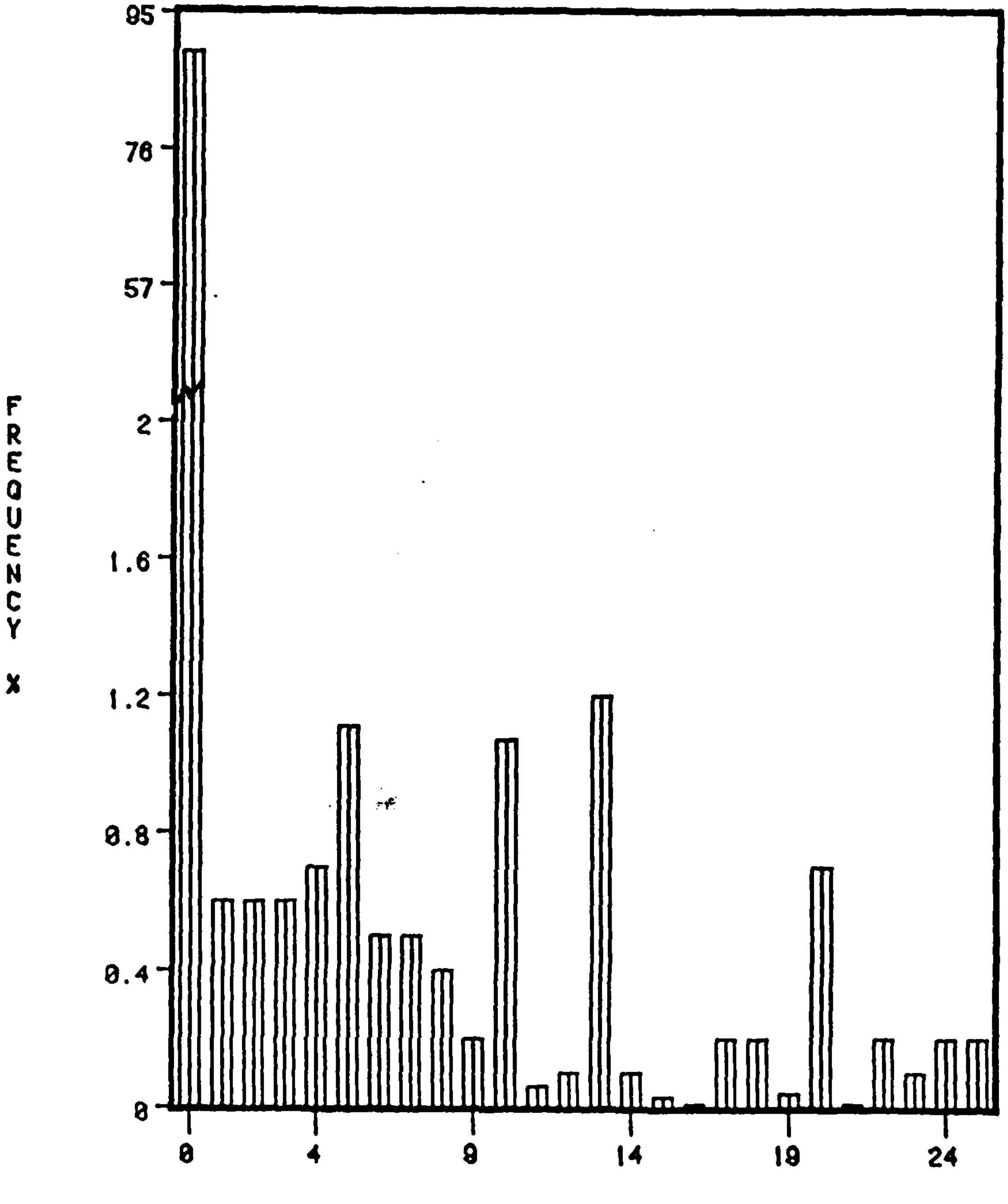


FIGURE 4: FREQUENCY OF POINTS USED FOR THE 25-POINT SCALE

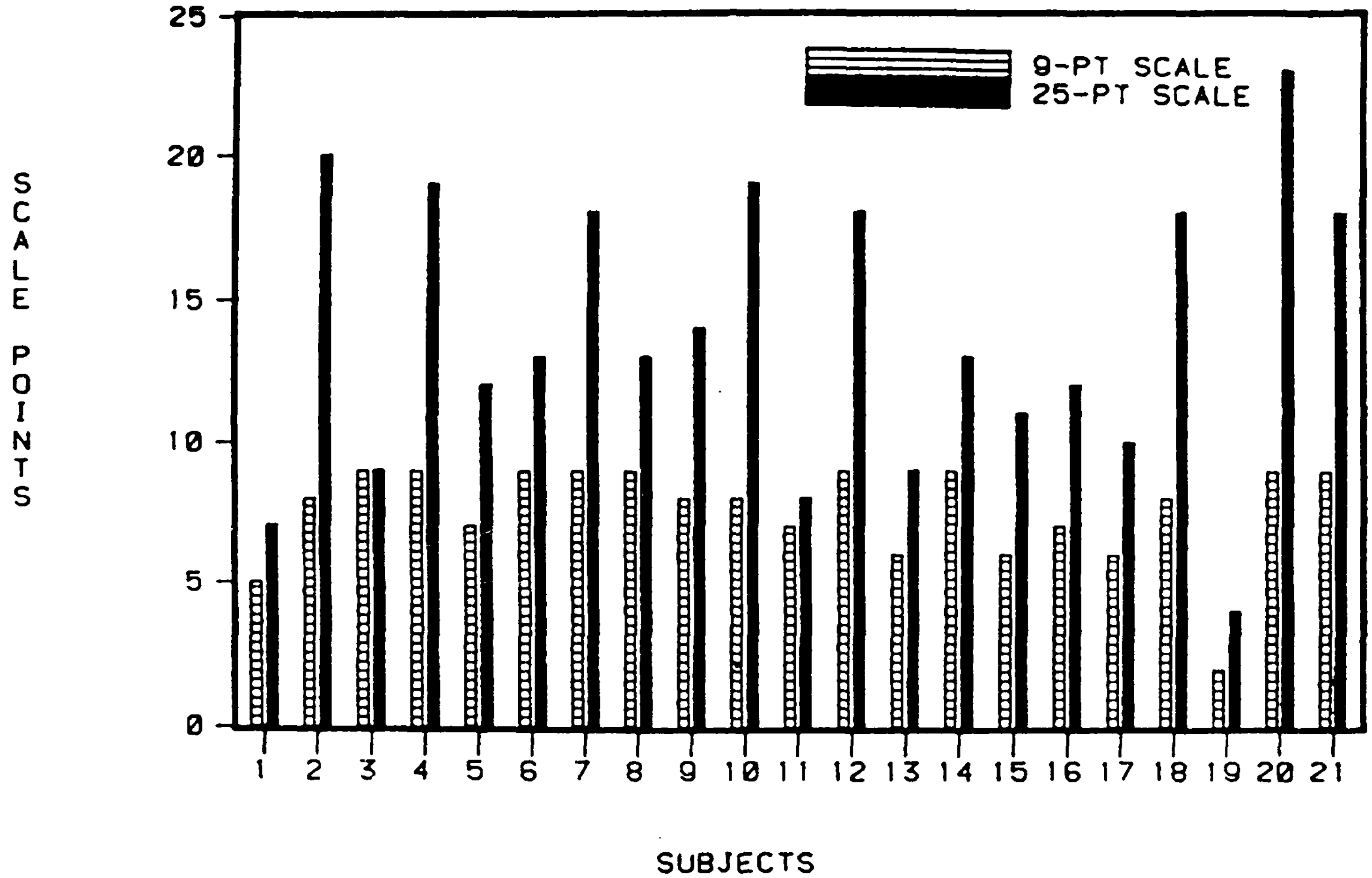


FIGURE 5: NUMBER OF SCALE POINTS USED BY SUBJECT

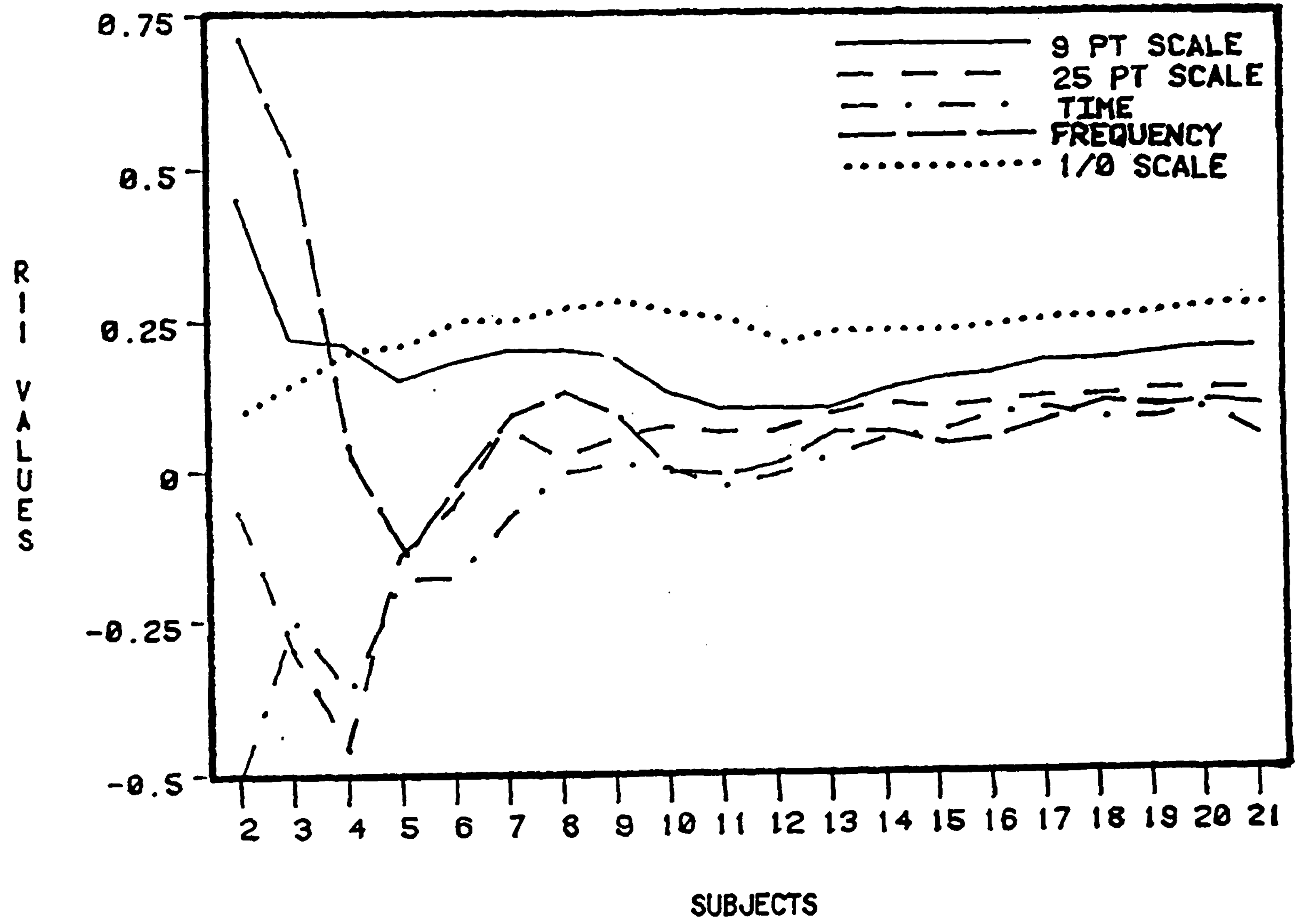


FIGURE 6: STABILITY OF VARIABLES (RII) AT DIFFERENT SAMPLE SIZES

25-point scales, with the percent of overlap being 86%. The perform/not perform scale compared highly with the 9-point scale (83%) and 25-point scale (82%) but dropped considerably with time (42%) and frequency (45%). Of the three predictors, the 9-point scale had the highest overlap with time (48%) and frequency (52%), followed by the 25-point scale. However, the difference in overlap between the 9- and 25-point scales was not found to be significant. The criterion of time was found to overlap highly with the criterion of frequency (75%). Additional information on duty and task-level comparisons for each scale is presented in the Appendices.

Table 5. Degree of Overlap Between Predictor/Criterion Variables

VARIABLES	1	2	3	4	5
1. PERFORM/NOT PERFORM SCALE	100	83	82	42	45
2. 9-POINT SCALE		100	86	48	52
3. 25-POINT SCALE			100	46	51
4. TIME				100	75
5. FREQUENCY					100

NOTE: These values are represented as percentages. Boxed-in area highlights relationships between predictors (1-3) and criteria (4-5).

Intercorrelations among the three predictors and two criteria are presented in Table 6. The 9-point scale correlated highly with the 25-point (r = .97) and perform/not perform scales (r = .94) but correlated somewhat lower with actual time (.79) and frequency (.81), though the values obtained still remain significant ($p \leq .05$). The 25-point scale correlated highly with the perform/not perform scale (.94) but not as high with time (.77) and frequency (.78). Correlation values obtained for the 1/0 scale and the two criterion variables were also found to be significant ($p \leq .05$). The correlation values for the binary scale dropped to .66 for time and .68 for frequency. A correlation value of .99 between the criteria time and observed frequency, indicated a high linear relationship between the two variables.

Table 6. Intercorrelations Among Predictor/Criterion Variables

VARIABLE	1	2	3	4	5
1. PERFORM/NOT PERFORM SCALE	1.00	.94	.94	.66	.68
2. 9-POINT SCALE		1.00	.97	.79	.81
3. 25-POINT SCALE			1.00	.77	.78
4. TIME				1.00	.99
5. FREQUENCY					1.00

Note: Boxed-in area highlights relationships between predictor (1-3) and criteria (4-5).

VI. DISCUSSION

The analyses reported herein used a direct observational approach as a criterion for validating task-level time spent rating scales. Previous research to validate such scales used other criteria, such as supervisory ratings of incumbent's time spent (Carpenter et al., 1975), number of scale points used (Wissman, 1980), reliability of scales (Pass & Robertson, 1978) and various other indirect techniques. Although these studies have resulted in some interesting findings regarding such factors as rating tendencies and the accuracy of derived job descriptions utilizing various scale formats, their results nevertheless are somewhat inconsistent and, therefore, inconclusive. On the other hand, research utilizing objective methods for the evaluation of self-reports, such as the observational approach, has provided greater consistency and conclusiveness in assessing the accuracy of self-reports. In two specific studies utilizing a direct observational approach, it was concluded that (a) on-site observation is an appropriate criterion for validating self-reports (Hartley et al., 1977) and (b) that the data collected by job inventories were found to be reliably related to the observation criteria (Johnson et al., 1980). Results of the present investigation support both these conclusions; that is, a direct observational approach and, in particular, the use of absolute time measures serve as an appropriate method to use in the assessment of the validity of relative-time spent-rating scales.

In evaluating the number of categories that can validly be used in estimating time spent on tasks, the analyses indicated that the accuracy of the job descriptions is non-linearly related to the number of response options available to the respondent. Accuracy did not decline as the number of response options decreased from 25 to 9. A decrease in accuracy was found,

however, when the options were reduced to 2. These findings support earlier results by Carpenter and Giorgia (1979), that an increase in time spent variance as a function of increasing the number of scale points is not significant, but that the potential inaccuracy of job descriptions resulting from fewer scale value options available is significant.

The average number of points used across all subjects for the 9-point scale was found to be seven as opposed to six found by Wissman (1980). For the 25-point scale, 13 was found to be the average number of points used across subjects, again in opposition to Wissman's findings of seven. In the present effort, it was determined that more scale points were being utilized than previously discussed in the literature. This discrepancy may be a function of the difference between the Fuels specialty studied by Wissman and the Computer Programming specialty studied here.

Of the three scales, the perform/not perform scale was found to be the least valid indicator of time spent when correlated with the two observational criteria. Results support previous research findings by Carpenter and Giorgia (1979), that the 9-point scale led to greater accuracy in estimating actual time spent in the performance of tasks. Although it has been argued that the use of a perform/not perform scale greatly reduces the amount of time required for completing a job inventory (Pass & Robertson, 1978), the suitability of this approach in describing jobs is greatly restricted by the limited number of response options available to the respondent.

When compared with time and frequency, the 9-point scale was the most valid measure to use in deriving relative time spent estimates of actual tasks performed. The 25-point scale was also found to be a highly valid predictor of actual time when compared with time and frequency, though slightly lower than for the 9-point scale.

The degree of overlap between predictor and criterion variables on the basis of tasks performed and relative time spent indicated that a greater degree of overlap exists between the 9-point scale and the two criteria than between the other two scales and the criteria. This high overlap lends significant support to the continued use of the 9-point scale in occupational measurement programs. The 25-point scale also overlapped highly with time and frequency, though slightly lower than the 9-point. An overlap of less than 50% was found with the perform/not perform scale when compared to the criteria.

The sample size of 21 was more than sufficient for this effort. The results indicated that with a sample of 15 or more, all five scales had little or no change in reliability with an increase in the number of cases. An advantage of using small samples is that it can allow for a closer observation of possible variables intervening to effect the outcome.

Perhaps the greatest limitation involved in the use of an observational approach is that such a method must be systematically controlled so as to minimize the effects of intervening variables (e.g., observer-subject interaction effects and time-ordered effects). If control is not incorporated within the design and maintained throughout, these variables could contaminate

the results. In this investigation, order effects arising from inventory administration, experimenter-subject effects and time of task performance effects were the main variables that were controlled. Control was achieved through the counterbalancing of inventory administration, alternating subjects across observers and varying observation times. Further, it allowed for a greater assessment of an observational approach as a small sample means of validating relative-time-spent rating scales.

VII. CONCLUSIONS

As a result of this effort, it was concluded that (a) the 9-point relative-time-spent scale provides the most accurate format for use in the Air Force occupational analysis program, (b) no significant gain in accuracy of information is found between 9- and 25-point scales, (c) the average number of points used was seven for the 9-point scale and 13 for the 25-point scale, (d) the perform/not perform scale was found to be the least accurate method of collecting time spent information, and (e) the use of an observational approach to yield absolute time estimates, and in particular, time and frequency, is appropriate for validating job descriptive scales.

References

Archer, W. B. (1966). Computation of group job description from occupational survey data (PRL-TR-66-12, AD-653 543). Lackland AFB, TX: Personnel Research Laboratory.

Carpenter, J. B., Giorgia, M. J., & McFarland, B. P. (1975). Comparative analysis of the relative validity for subjective time rating scales (AFHRL-TR-75-63, AD-A017 842). Lackland AFB, TX: Occupational and Manpower Research Division, Air Force Human Resources Laboratory.

Carpenter, J. B., & Giorgia, M. J. (1979). Relative time-spent-rating scales - another look at performance measures. Proceedings, 21st annual conference of the Military Testing Association, San Diego, CA.

Christal, R. E. (1974). The United States Air Force occupational research project (AFHRL-TR-73-75, AD-774 575). Lackland AFB, TX: Occupational Research Division, Air Force Human Resources Laboratory.

Christal, R. E., & Madden, J. (1961). Air Force research on job evaluation procedures (ASD-TN-61-46, AD-267 346) Lackland AFB, TX: Personnel Research Laboratory.

Christal, R. E., & Weissmuller, J. J. (1976). New CODAP programs for analyzing task factor information (AFHRL-TR-76-3, AD-A026 121). Lackland AFB, TX: Occupational and Manpower Research Division, Air Force Human Resources Laboratory.

Hartley, C., Brecht, M., Pagery, P., Weeks, G., Chapanis, A., & Hoecker, D. (1977). Subjective time estimates of work tasks by office workers. Journal of Occupational Psychology, 50, 23-26.

Johnson, C. A., Hiller, J., & Tokunaga, H. T. (1980). Validation of A Job Analysis Questionnaire Through Intensive Observation. Proceedings, 22nd annual conference of the the Military Testing Association, Toronto, Canada.

McCormick, E. J. Job and task analysis. In M. D. Dunnette (Ed). (1976). Handbook of Industrial and Organizational Psychology. Chicago: Rand McNally College Publishing Company, 651-696.

Pass, J. J., & Robertson, D. W. (1978). Sample Size and Stability of Task Analysis Inventory Response Scales. Proceedings, 20th annual conference of the Military Testing Association, (pp. 537-555). Oklahoma City, OK.

Wissman, D. (1980). Design and validation of relative time spent rating scales Paper presented at the 3d annual Conference of Military Occupational Analysts, Randolph Air Force Base, TX.

Appendix A : Comparisons Between Scales on Percent Time Spent by Duty and Task Level

Table A-1. Comparison of Percent Time Spent on Duties for Criteria and Predictors

DUTY	TIME	FREQUENCY	1/0	9PT	25PT
D PROGRAMMING FUNCTION	69	64	37	44	41
E ANALYSIS FUNCTION	10	9	13	13	14
B PRODUCTION CONTROL FUNCTION	9	13	17	17	19
A OPERATOR FUNCTION	3	4	9	8	9
G TRAINING FUNCTION	3	3	5	4	5
H SUPERVISORY FUNCTION	2	3	6	4	4
I ADMINISTRATIVE FUNCTION	2	2	3	2	2
C LIBRARY FUNCTION	1	1	1	1	1
J SUPPLY & CONTRACTING FUNCTION	1	1	1	1	1
F SECURITY FUNCTION	*	1	5	4	3
K GENERAL MILITARY FUNCTION	*	*	3	2	1
TOTAL:	100	100	100	100	100

*DENOTES LESS THAN 1 PERCENT TIME SPENT

Table A-2. Representative Tasks Performed Using the 9- and 25-PT, 1/0, Time and Frequency Scales

PERCENT OF TIME SPENT

TASKS	9PT	25PT	1/0	Time	Freq
D 176 DEBUG COMPUTER PROGRAMS	4	5	3	15	14
D 179 DESK CHECK PROGRAMS	4	4	3	9	8
D 162 CODE COMPUTER PROGRAMS IN HIGHER LEVEL LANGUAGES	4	3	3	6	5
D 165 CODE DATA SET UTILITY PROGRAMS	4	3	2	3	3
D 186 MODIFY OR UPDATE EXISTING COMPUTER PROGRAMS	3	4	3	10	9
D 175 COMPILE OR ASSEMBLE PROGRAMS	3	2	2	3	3
D 199 TEST COMPUTER PROGRAMS	3	2	2	4	4
D 180 DETERMINE CAUSES OF PROGRAM HALTS OR ABENDS	3	4	2	*	*
D 159 ALLOCATE IMMEDIATE ACCESS STORAGE	3	2	3	*	*
D 190 PREPARE DETAILED FLOWCHARTS	3	3	3	2	2
B 105 BATCH RUN REQUEST	3	2	2	*	*
B 104 ASSEMBLE, REARRANGE, OR EDIT INPUT OR OUTPUT DATA	2	3	2	*	*
B 103 ANALYZE OUTPUT PRODUCTS FOR COMPLIANCE WITH STANDARDS OR SPECIFICATIONS		2	3	2	*
B 111 DETERMINE CAUSE OF FAULTY OUTPUT PRODUCTS	2	3	2	*	1
D 191 PREPARE DOCUMENTATION FOR INDIVIDUAL PROGRAMS	2	2	2	1	1
B 133 RESPOND TO INQUIRIES FROM CUSTOMERS	2	2	2	*	*
B 113 DISTRIBUTE OR DELIVER OUTPUT PRODUCTS	2	2	2	1	4
F 334 STORE OR SAFEGUARD PRIVACY ACT INFORMATION	2	2	2	*	*
B 110 DETERMINE ALTERNATE METHODS FOR ACCOMPLISHING JOB REQUIREMENTS	2	2	2	*	*
D 171 CODE SERVICE AID UTILITY PROGRAMS	2	*	*	*	*
D 198 REVIEW PROGRAM SPECIFICATIONS	2	2	2	*	*
D 184 EXPLAIN ERROR PRINTOUTS TO CUSTOMERS	2	2	2	*	*
D 197 REVIEW PROGRAM REQUIREMENTS	1	2	2	2	3
D 181 DETERMINE PROGRAM RUN TIMES	1	2	2	*	*
F 318 MARK PRIVACY ACT INFORMATION	1	1	2	*	*
D 182 DEVELOP MODELS OR DUMMY DATA TO SIMULATE FUNCTIONAL REQUIREMENTS	1	2	2	*	*

Note: An asterisk denotes less than 1 percent of time spent

Table A-3. Tasks Which Best Distinguish Between Actual Time and 1/0 Scale Based on Percent Time Spent

DUTY/TASK	TIME	1/0	DIFFERENCE
D 176 DEBUG COMPUTER PROGRAMS	12.18	2.41	+9.77
D 186 MODIFY OR UPDATE EXISTING COMPUTER PROGRAMS	7.75	2.31	+5.44
D 179 DESK CHECK PROGRAMS	7.45	2.31	+5.14
D 161 CODE COMPUTER PROGRAMS IN ASSEMBLY LANGUAGES	3.58	.85	+2.73
D 162 CODE COMPUTER PROGRAMS IN HIGHER LEVEL LANGUAGES	4.73	2.26	+2.47
E 225 DETERMINE INTERRELATIONSHIPS AMONG FILES, DOCUMENTS, AND ITEMS	2.46	.71	+1.75
D 199 TEST COMPUTER PROGRAMS	3.42	1.75	+1.67
D 197 REVIEW PROGRAM REQUIREMENTS	1.85	.94	+ .91
D 165 CODE DATA SET UTILITY PROGRAMS	2.23	1.36	+ .87
D 166 CODE FOR GRAPHIC DISPLAY PLOTTERS	.57	.04	+ .53
D 193 PREPARE PROGRAM TEST SPECIFICATIONS OR INSTRUCTIONS	.80	.28	+ .52
D 198 REVIEW PROGRAM SPECIFICATIONS	.24	1.24	-1.01
D 184 EXPLAIN ERROR PRINTOUTS TO CUSTOMERS	.04	1.09	-1.05
B 105 BATCH RUN REQUEST	.01	1.10	-1.09
D 180 DETERMINE CAUSES OF PROGRAM HALTS OR ABENDS	.71	1.96	-1.25
F 318 MARK PRIVACY ACT INFORMATION	.00	1.35	-1.35
F 334 STORE OR SAFEGUARD PRIVACY ACT INFORMATION	.17	1.53	-1.35
D 181 DETERMINE PROGRAM RUN TIMES	.00	1.50	-1.50
D 159 ALLOCATE IMMEDIATE ACCESS STORAGE	.13	1.84	-1.72

Table A-4. Tasks Which Best Distinguish Between Actual Time and 9-Point Scale Based on Percent Time Spent

DUTY/TASK	TIME	9PT	DIFFERENCE
D 176 DEBUG COMPUTER PROGRAMS	12.18	4.08	+8.11
D 186 MODIFY OR UPDATE EXISTING COMPUTER PROGRAMS	7.75	2.95	+4.81
D 179 DESK CHECK PROGRAMS	7.45	4.05	+3.40
D 161 CODE COMPUTER PROGRAMS IN ASSEMBLY LANGUAGES	3.58	1.09	+2.50
E 225 DETERMINE INTERRELATIONSHIPS AMONG FILES, DOCUMENTS AND ITEMS	2.46	.59	+1.89
D 162 CODE COMPUTER PROGRAMS IN HIGHER LEVEL LANGUAGES	4.73	3.11	+1.63
D 199 TEST COMPUTER PROGRAMS	3.42	2.20	+1.22
D 197 REVIEW PROGRAM REQUIREMENTS	1.85	.89	+ .96
D 166 CODE FOR GRAPHIC DISPLAY PLOTTERS	.57	.03	+ .54
D 181 DETERMINE PROGRAM RUN TIMES	.00	1.01	-1.01
D 167 CODE INSTRUCTIONS TO INTERCEPT END OF JOBS (ABENDS)	.00	1.10	-1.10
D 180 DETERMINE CAUSES OF PROGRAMS HALTS OR ABENDS	.71	2.11	-1.40
B 105 BATCH RUN REQUESTS	.01	1.52	-1.51
D 159 ALLOCATE IMMEDIATE ACCESS STORAGE	.13	1.90	-1.78

Table A-5. Tasks Which Best Distinguish Between Actual Time and 25-Point Scale Based on Percent Time Spent

DUTY/TASK	TIME	25PT	DIFFERENCE
D 176 DEBUG COMPUTER PROGRAMS	12.18	3.70	+8.49
D 186 MODIFY OR UPDATE EXISTING COMPUTER PORGRAMS	7.75	3.21	+4.54
D 179 DESK CHECK PROGRAMS	7.45	4.25	+3.20
D 161 CODE COMPUTER PROGRAMS IN ASSEMBLY LANGUAGES	3.58	.89	+2.69
D 162 CODE COMPUTER PROGRAMS IN HIGHER LEVEL LANGUAGES	4.73	3.04	+1.70
D 199 TEST COMPUTER PROGRAMS	3.42	1.74	+1.68
E 225 DETERMINE INTERRELATIONSHIPS AMONG FILES, DOCUMENTS, AND ITEMS	2.46	1.11	+1.34
B 105 BATCH RUN REQUESTS	.01	1.20	-1.20
D 181 DETERMINE PROGRAM RUN TIMES	.00	1.25	-1.25
D 159 ALLOCATE IMMEDIATE ACCESS STORAGE	.13	1.54	-1.41
B 111 DETERMINE CAUSE OF FAULTY OUTPUT PRODUCTS	.77	2.21	-1.44
D 180 DETERMINE CAUSES OF FAULTY HALTS OR ABENDS	.71	2.75	-2.04

Table A-6. Tasks Which Best Distinguish Between Relative Frequency and 1/0 Scale Based on Percent Time Spent

DUTY/TASK	FREQUENCY	1/0	DIFFERENCE
D 176 DEBUG COMPUTER PROGRAMS	11.91	2.41	+9.50
D 186 MODIFY OR UPDATE EXISTING COMPUTER PROGRAMS	7.72	2.31	+5.41
D 179 DESK CHECK PROGRAMS	7.43	2.31	+5.12
B 113 DISTRIBUTE OR DELIVER OUTPUT PRODUCTS	3.48	1.25	+2.23
D 161 CODE COMPUTER PROGRAMS IN ASSEMBLY LANGUAGES	2.93	.85	+2.07
D 162 CODE COMPUTER PROGRAMS IN HIGHER LEVEL LANGUAGES	4.31	2.26	+2.04
E 225 DETERMINE INTERRELATIONSHIPS AMONG FILES, DOCUMENTS, AND ITEMS	2.37	.71	+1.66
D 199 TEST COMPUTER PROGRAMS	3.34	1.75	+1.59
D 197 REVIEW PROGRAM REQUIREMENTS	2.36	.94	+1.43
D 165 CODE DATA SET UTILITY PROGRAMS	2.37	1.36	+1.01
D 184 EXPLAIN ERROR PRINTOUTS TO CUSTOMERS	.08	1.09	-1.01
B 105 BATCH RUN REQUESTS	.07	1.10	-1.03
D 180 DETERMINE CAUSES OF PROGRAM HALTS OR ABENDS	.64	1.96	-1.32
F 318 MARK PRIVACY ACT INFORMATION	.00	1.35	-1.35
F 334 STORE OR SAFEGUARD PRIVACY ACT INFORMATION	.16	1.53	-1.37
D 181 DETERMINE PROGRAM RUN TIMES	.00	1.50	-1.50
D 159 ALLOCATE IMMEDIATE ACCESS STORAGE	.24	1.84	-1.60

Table A-7. Tasks Which Best Distinguish Between Relative Frequency and 9-Point Scale Based on Percent Time Spent

DUTY/TASK	FREQUENCY	9PT	DIFFERENCE
176 DEBUG COMPUTER PROGRAMS	11.91	4.08	+7.83
D186 MODIFY OR UPDATE EXISTING COMPUTER PROGRAMS	7.72	2.95	+4.77
D179 DESK CHECK PROGRAMS	7.43	4.05	+3.38
B113 DISTRIBUTE OR DELIVER OUTPUT PRODUCTS	3.48	1.14	+2.34
D161 CODE COMPUTER PROGRAMS IN ASSEMBLY LANGUAGES	2.93	1.09	+1.84
E225 DETERMINE INTERRELATIONSHIPS AMONG FILES, DOCUMENTS, AND ITEMS	2.37	.57	+1.80
D197 REVIEW PROGRAM REQUIREMENTS	2.36	.89	+1.48
D162 CODE COMPUTER PROGRAMS IN HIGHER LEVEL LANGUAGES	4.31	3.11	+1.20
D199 TEST COMPUTER PROGRAMS	3.34	2.20	+1.13
D181 DETERMINE PROGRAM RUN TIMES	.00	1.01	-1.01
D167 CODE INSTRUCTIONS TO INTERCEPT ABNORMAL END OF JOBS (ABENDS)	.00	1.10	-1.10
B105 BATCH RUN REQUESTS	.07	1.52	-1.45
D180 DETERMINE CAUSES OF PROGRAM HALTS OR ABENDS	.64	2.11	-1.47
D159 ALLOCATE IMMEDIATE ACCESS STORAGE	.24	1.90	-1.66

Table A-8. Tasks Which Best Distinguish Between Relative Frequency and 25-Point Scales Based on Percent Time Spent

DUTY/TASK	FREQUENCY	25PT	DIFFERENCE
D 176 DEBUG COMPUTER PROGRAMS	11.91	3.70	+8.22
D 186 MODIFY OR UPDATE EXISTING COMPUTER PROGRAMS	7.72	3.21	+4.51
D 179 DESK CHECK PROGRAMS	7.43	4.25	+3.18
B 113 DISTRIBUTE OR DELIVER OUTPUT PRODUCTS	3.48	.91	+2.57
D 161 CODE COMPUTER PROGRAMS IN ASSEMBLY LANGUAGES	2.93	.89	+2.04
D 199 TEST COMPUTER PROGRAMS	3.34	1.74	+1.60
D 197 REVIEW PROGRAM REQUIREMENTS	2.36	1.06	+1.30
D 162 CODE COMPUTER PROGRAMS IN HIGHER LEVEL LANGUAGES	4.31	3.04	+1.27
E 225 DETERMINE INTERRELATIONSHIPS AMONG FILES, DOCUMENTS, AND ITEMS	2.37	1.11	+1.25
F 334 STORE OR SAFEGUARD PRIVACY ACT INFORMATION	.16	1.16	-1.01
B 105 BATCH RUN REQUESTS	.07	1.20	-1.13
D 181 DETERMINE PROGRAM RUN TIMES	.00	1.25	-1.25
D 159 ALLOCATE IMMEDIATE ACCESS STORAGE	.24	1.54	-1.30
B 111 DETERMINE CAUSES OF FAULTY OUTPUT PRODUCTS	.89	2.21	-1.33
D 180 DETERMINE CAUSES OF PROGRAM HALTS OR ABENDS	.64	2.75	-2.11

END

FILMED

9-84

DTIC